



Hydrodynamic backtracking of fish larvae by individual-based modelling

Asbjørn Christensen^{1,*}, Ute Daewel², Henrik Jensen¹, Henrik Mosegaard¹,
Mike St. John², Corinna Schrum^{1,3}

¹Danish Institute of Fisheries Research (DIFRES), Charlottenlund Slot, 2920 Charlottenlund, Denmark

²Institute for Hydrobiology und Fisheries Science, Hamburg University, Olbersweg 24, 22767 Hamburg, Germany

³Geophysical Institute, University of Bergen, Allégaten 70, 5007 Bergen, Norway

ABSTRACT: We discuss methodological and implementation issues of spatial, temporal and combined spatio-temporal backtracking and illustrate larval backtracking for North Sea lesser sandeel *Ammodytes marinus* larvae, using a combined hydrodynamical and individual-based model. It was found that dispersal effects are important for larval backtracking predictions. Our results show large differences in average transport distance, as well as in shape and extent of predicted hatch areas, when backtracking advected larval cohorts in different regions of the North Sea, thus emphasizing the importance of using realistic, spatially and temporally resolved diffusivity fields in simulations of larval transport. In all cases, biologically likely hatching areas have been predicted. We discuss issues of methodological consistency and present a new scheme for including life-history stochasticity effects on growth in backtracking in a consistent way, as well as procedures for assessing the effects of larval mortality. Finally, fundamental limitations of larval backtracking are clarified, most importantly the time horizon and spatial resolution limit for backward prediction.

KEY WORDS: Individual-based modelling · Backtracking · Larval transport · Inverse stochastic methods · Hatch area identification

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Backtracking fish larvae is a potentially powerful tool for understanding early life-history aspects quantitatively. Backtracking may be performed either explicitly in space and time or just in time—in the latter case with implicit spatial assumptions about the environmental parameters affecting larval growth and behaviour. The goal of spatial backtracking is mapping hatching areas with high resolution, to understand conditions for survival (or larval excess mortality) and thereby develop minimal conceptual models of recruitment, based on key processes. Also, optimizing the design of closed marine habitats with respect to habitat definition and closure time is a potential application of backtracking.

Backtracking can also be used as an efficient biological parameter estimation tool, which may supplement

forward tracking of larvae; the biological parameters of a larval tracking model should be fitted so that a simulation develops larvae into juveniles at the right area and time. If our knowledge on juveniles is more precise than on newly hatched larvae, it may require less trajectory simulations to backtrack from juveniles (because more trajectories will be successful). Additionally, biological parameter sensitivity may be different in backtracking, compared to forward tracking.

While backtracking is frequently used in other fields like pure computer sciences (Dechter & Frost 2002), atmospheric sciences (Uliasz & Pielke 1991) and pollution tracing (Spivakovskaya et al. 2005), larval/zoo-plankton backtracking is still at a premature stage, seen from a methodological point of view. However, recently Batchelder (2006) applied a backtracking scheme to planktonic organisms in a coastal geometry and advocated backtracking as a tool for identifying

*Email: asc@difres.dk

hatching sites. We go one step further and consider growth backtracking as well in a fully realistic hydrodynamic setup. Compared to passive particles, larvae have ontogenetic development as well as active behaviour, which complicates backtracking, due to the stochastic and nonlinear nature of these processes.

At first sight, the issue of tracing backwards the state of a larval sample in time appears simple: just advect larvae in a direction opposite the currents and shrink their size by the same amount as their growth, if time is running forward (Pedersen et al. 2000, Allain 2004); however, dispersion processes are present in the ocean and must be considered, as we will show later, because otherwise we have no information on the error margin in results and further simple backtracking opposite current lines may contain an error bias, since dispersion processes in the ocean are not spatially uniform. Batchelder (2006) similarly found that diffusion cannot be ignored in backtracking. Alternatively, larval origin might be estimated by a forward approach by releasing a vast number of particles in a forward simulation at potential spawning grounds and focussing on the small number that arrives in the area of interest (Allain et al. 2003). While it is straightforward and circumvents some of the methodological issues of backtracking, forward tracking remains inefficient for spawning-ground identification because the majority of particle trajectories are useless (they end up at a position different from the area of interest). This is especially true if arrival in the area of interest is a rare event (e.g. if the final area is small or the drift time is long); this point was also emphasized by Batchelder (2006). Further, backtracking may potentially point to unexpected spawning sites, whereas forward tracking has less room for surprises because potential spawning grounds are input to the simulation.

The aim of the present paper is to advocate larval backtracking as part of a testing suite of hydrodynamical individual-based models (IBM) for larvae, both as a submodel screening and validation device, as well as a result-generating method, and to clarify the formal basis of backtracking with focus on providing tools for consistent backtracking and identifying limitations of backtracking. The latter aspect is very important, because inverse problems often have no unique solution.

BIOLOGICAL AND PHYSICAL MODEL

In Lagrangian transport simulations, the positions ($s_i[t]$) of an ensemble of tracers $i = 1 \dots N$ are monitored as a function of time t , along with the state variables ($L_i[t]$) of the tracers. This ensemble represents the fish larvae, with their length L_i as state variable, in our

case. Each larva in the ensemble is propagated from time t to $t + dt$ by the dynamical equations:

$$dL_i = G(L_i, s_i, t)dt \quad (1)$$

$$ds_i = [u(s_i, t) + a(s_i, L_i, t)]dt + d\Omega(s_i, t, dt) \quad (2)$$

$$d\Omega(s_i, t, dt) = \nabla K(s_i, t)dt + \sqrt{2K[s_i + \frac{1}{2}\nabla K(s_i, t)dt, t]}W(dt) \quad (3)$$

where dt is a small positive time increment. Eq. (1) describes the larval growth $L_i \rightarrow L_i + dL_i$, which depends on larval size, position and time in season through a deterministic relation G . It is also possible to model G as a stochastic relation; we will return to this in section 'Growth stochasticity'. Eq. (2) describes the larval transport $s_i \rightarrow s_i + ds_i$, where $u(s_i, t)$ is the smooth advective field derived by linear interpolation in grid-resolved currents, obtained from the hydrodynamical model described below, and $d\Omega$ is a random walk process modelling the effect of local turbulent fluctuations (Visser 1997). The active larval motion velocity is $a(s_i, L_i, t)$, including possibly buoyant velocity. In Eq. (3) $d\Omega$ is controlled by the local diffusivity field $K(s_i, t)$, and $W(dt)$ is a stochastic process with $\langle W(dt) \rangle = 0$ and $\langle W(dt)^2 \rangle = dt$, where $\langle \rangle$ indicates time average. Because of this, the limit $d\Omega / dt$ does not exist, and Eqs. (1) to (3) cannot be stated as ordinary partial differential equations, but are kept in Itô form (Rogers & Williams 1987) as above. In Cartesian coordinates, the local diffusivity field is a vector $K = (K_x, K_y, K_z)$, and the gradient diagonal is $\nabla K = (dK_x/dx, dK_y/dy, dK_z/dz)$. The spatial distribution of an ensemble of larvae moving by Eq. (2) can also be described by the physics diffusion equation with local diffusivity K (Taylor 1921, Visser 1997), and equivalent results are obtained generally if consistent dynamical equations are used and the Lagrangian ensemble size is sufficiently large. The local diffusivity field K describes subgrid advective processes, like turbulence, in a statistical sense. Eqs. (1 to 3) describe virtual larval trajectories, i.e. without mortality. We return to mortality aspects in the section 'Discussion'.

Forward time model. A flexible, spatially explicit lesser sandeel *Ammodytes marinus* IBM has been developed on the basis of the ECOSMO (ECOSystem MOdel) framework (Schrum & Backhaus 1999, Hochbaum 2004, Schrum et al. 2006). The ECOSMO hydrodynamic model setup has been validated in detail against available observations (Janssen et al. 2001, Janssen 2002). In the present work, a biological model of the early larval life-stages of the lesser sandeel has been added, based on the underlying biophysical processes. The setup is shown in Fig. 1.

The hydrodynamic part of the ECOSMO model is based on a staggered Arakawa C-grid with a 5 nm horizontal resolution, free surface and 5 m layers down to a depth of 40 m (and 8 m layers below 40 m depth).

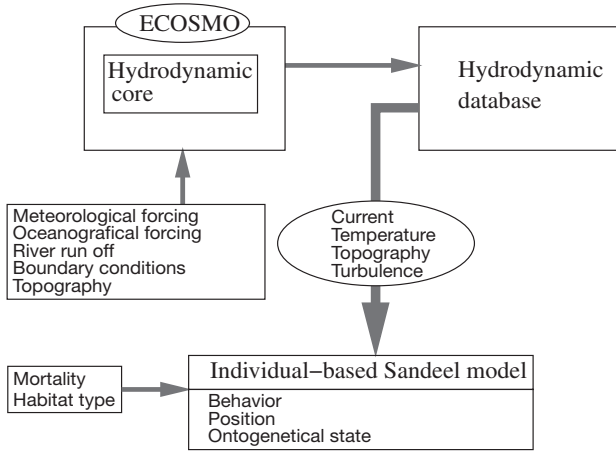


Fig. 1. Coupled 3-dimensional biophysical model. ECOSMO: ECOSystem Model

A database (Schrum et al. 2003) of 3-dimensional physical fields from the ECOSMO model—currents $u(s,t)$, temperature $T(s,t)$ and vertical local diffusivity $K_z(s,t)$ —has been stored as daily averaged fields (for data compression purposes), and these fields are used for the IBM simulations. $K_z(s,t)$ is parameterized using an analytical k - ϵ approach, considering counteracting effects of local shear and stratification (Schrum 1997).

In the present work, we include no explicit active vertical/horizontal migratory behaviour, i.e. $a = 0$ in Eq. (2). This is a good approximation during night time, but during light hours there is a predominance of larvae in the water layers with high zooplankton abundance (Jensen et al. 2003). However, until the vertical behaviour has been accurately quantified for lesser sandeel, we will use $a = 0$.

Horizontally, the larvae are described as passive floaters. Only vertical turbulent dispersal is taken into account, i.e. $K = (0, 0, K_z)$, and $d\Omega$ is a vertical vector, since the dominant horizontal dispersal mechanism in the North Sea is vertical turbulent diffusion, coupled to current layer shear (Zimmerman 1986, Van Dam et al. 1999). As larvae hatch and are advected, they disperse relative to each other. Within the dispersed patch—which is on a scale of kilometres—there may be small-scale patchiness, due to larval schooling behaviour and subscale environmental patchiness. We will not consider these additional small-scale variations in the larval spatial distribution in this work, but rather concentrate on the kilometre-scale features in transport and dispersal of larval patches from the same area.

The larval population is mathematically sampled by a set of representative tracers (a virtual population), each of which represents a constant number of individuals (the ratio of real physical larvae per tracer needs not be stipulated, since density effects are not addressed explicitly in the present study, only relative

numbers matter). The boundary condition $dK_z(s,t)/dz = 0$ along with tracer reflection is imposed vertically at the surface and bottom to avoid artificial aggregation of tracers at the surface or bottom of the water column. The vertical random walk of Eq. (3) is implemented as $W(dt) = \omega\sqrt{3dt}$, where ω is a uniform random distribution on $[-1, 1]$, corresponding to $\langle W(dt)^2 \rangle = dt$, so that local jump amplitudes reproduce local Eulerian field dispersal rates correctly, i.e. proportional to the square root of the local diffusivity $K(s,t)$ (Taylor 1921, Maier-Reimer 1973, Hunter et al. 1993).

The Lagrangian simulations were performed with a time step $dt = 30$ min, using Euler forward integration. In the section ‘Spatial and temporal backtracking’ we will show that higher order horizontal integration schemes change tracer trajectories negligibly for $dt = 30$ min, i.e. trajectories are appropriately integrated numerically with $dt = 30$ min, when using current fields averaged over tidal periods. Longer time steps in conjunction with higher order horizontal trajectory integration were not attempted, since this would imply large vertical jumps in the stochastic modelling of turbulent dispersal (Eq. 3).

The larval growth model in Eq. (1) is parameterized to the functional form:

$$G(L, T) = \lambda(T) \left(\frac{L}{L_0} \right)^\gamma \left(1 - \frac{L}{L_\infty} \right) \quad (4)$$

where $T = T(s,t)$ is the local temperature experienced by each larva. The data set used for parameterization is North Sea length-at-age samples for *Ammodytes marinus* obtained by MIK trawl data from the years 1995 and 1996 (pooled together) (Jensen 2001). Larval ages were obtained by otolith analysis (Jensen 2001). The temperature modulation $\lambda(T)$ is approximated by a quadratic polynomial:

$$\lambda(T) = a_0 + a_1 T + a_2 T^2 \quad (5)$$

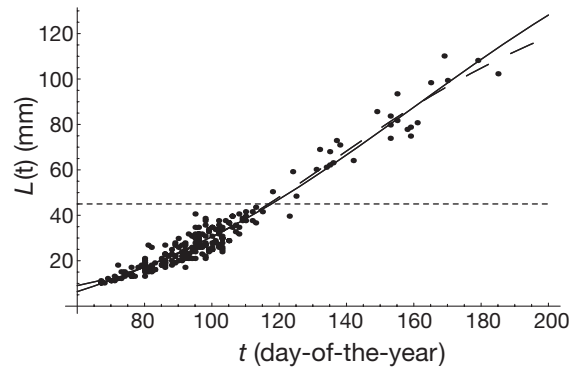


Fig. 2. Pooled length-age data (points) used for parameterizing the lesser sandeel *Ammodytes marinus* growth model. The horizontal line at $L_m = 45$ mm corresponds to the approximate length at metamorphosis. Continuous line: Model v1; dashed line: Model b1, see Table 1

In Fig. 2, we show the data and model fits, based on analytical integration of Eqs. (1) and (4), including the North Sea average seasonal temperature variation. It is important to include the average seasonal temperature variation in the parameter estimation process, since the average temperature rises from around 5°C at hatching time to around 12 to 14°C at the time when metamorphosis is complete ($L_m \approx 45$ mm) (Wright & Bailey 1996, Jensen 2001); otherwise, model forward simulation will not reproduce the growth pattern in Fig. 2. It appeared difficult to simultaneously resolve $\lambda(T)$ and the length scaling exponent γ , because the seasonal temperature variation and larval length are strongly correlated in the data set. There is a shallow residual minimum in the fit at $\gamma \approx 0.96$, but fits constrained to $\gamma = 0$ (Model v) or $\gamma = 1$ (Model b) produce essentially equally good fits for 0-group sandeels, as shown in Fig. 2. The parameters of the respective fits are given in Table 1, along with the estimated hatch lengths L_0 , all of which are consistent with observations (Winslade 1971, Smigielski et al. 1984) of $L_0 \approx 5$ to 7 mm. We use $L_\infty = 218$ mm (Macer 1966), which is not included as a free parameter when fitting data in Fig. 2, since it is of minor importance for larval growth (since $L_m \ll L_\infty$) and because the data set only has observations for $L < L_\infty$ (but L_∞ adds a little concavity to the growth curve).

Table 1. Parameters for the alternative sandeel *Ammodytes marinus* growth models. The fitting residual is normalized as $\rho = \sqrt{(1/M) \sum_{i=1}^M [L^u(t_i) - L_i]^2}$, where $(L_i, t_i)_{i=1, \dots, M}$ is the length-age data set and $L^u(t)$ is the growth curve for Model μ

Model μ	γ	L_0 (mm)	a_0 (mm d ⁻¹)	a_1 (mm d ⁻¹ °C ⁻¹)	a_2 (mm d ⁻¹ °C ⁻²)	ρ (mm)
v1	0	6.40	-0.354	0.167	0	10.8
v2	0	8.91	-1.81	0.515	-0.0196	11.9
b1	1	9.05	0.422	-0.0205	0	12.0
b2	1	9.15	0.401	-0.0153	-0.000272	12.1

From Fig. 2, it is seen that seasonal temperature variation is a surprisingly good proxy for the complex bioenergetic effects (spectrum, abundance and distribution of food, as well as complex food-switching patterns) (Letcher et al. 1996, Baron 2004). The quantitative growth predicted by Eq. (4), as parameterized here, is in good quantitative agreement with that presented recently by Gallego et al. (2004). Since the model parameterizes *in situ* data, it also integrates prey dependence on temperature variations, at a crude level. Currently, the model does not address density effects (e.g. food competition and cannibalism) explicitly, although they are believed to affect recruitment (Daan et al. 1990, Kishi et al. 1991, Kimura et al. 1992, Arnott & Ruxton 2002). There are also indications that population density effects may affect individual sandeel growth, currently reported for sandeel populations in the North Sea (Bergstad et al. 2002) and Ise Bay (Nagoshi & Sano 1979), but the influence has not yet been sufficiently quantified for larval stages.

Reversed time model. Let us start at the heart of the problem, which is shown in Fig. 3. Fig. 3a illustrates the familiar forward time modelling situation, where a larval patch (grey area) is traced from the hatching area to Patch J at a later time (which could symbolize the distribution of juvenile larvae). The larval patch increases in spatial extent, as the patch is advected along current lines, due to spatial dispersive processes. If we now catch a portion of the larvae in Patch J and consult the model as to where they came from, many people anticipate the situation in Fig. 3b: the model, when ‘run backwards’, should converge to the hatch area, where the forward simulation started from. We call this *inverse time* simulation. Since the Fourier spectrum of inverse time dispersion diverges for small spatial scales, the inverse time simulation is not numerically stable (Hadamard 1923). The physical meaning of this is that forward time diffusion quickly dissipates fine-scale features, while inverse time simulations enhances them. We want to calculate the initial distribution leading to a given final distribution

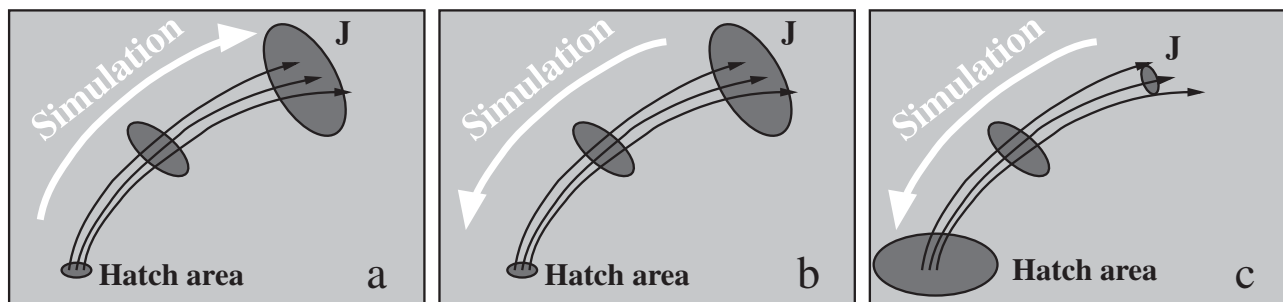


Fig. 3. Different types of hydrodynamical simulations. Black arrows: current stream lines. (a) Normal forward time simulation of a larval patch (grey area), starting from the hatch area with larvae and ending with a juvenile distribution at J. (b) Inverse time simulation of a larval patch (grey area), starting from a juvenile sample at J. (c) Reversed time simulation of a larval patch (grey area), starting from a juvenile sample at J

of larvae in a statistical sense; this is called *reverse time* simulation (Fig. 3c). The major difference, compared to inverse time simulation in Fig. 3b, is that the larval patch will disperse in space when traced backwards in time. The point is that dispersive (and generally stochastic) processes delete knowledge on initial state when time progresses, and therefore the uncertainty of the larval origin also increases when they are traced back in time. In other words, given a larva in a specific place and state, we can only give a spatial probability distribution of places where it is likely to have hatched because it can end up in a specific place and state along many different life-history paths, when dispersive (and generally stochastic) effects are present in the model. This limitation is fundamental and insurmountable, and has important consequences. First, if we neglect spatial dispersive processes (if this is possible) and just backtrack along current lines, we will end up at a point somewhere in the hatch area distribution in Fig. 3c, but we have no guarantee that the point is at the centre of the hatch area distribution (and it will not be at the centre, when the dispersive processes have spatial gradients, which is usually the case), and we have no idea about the characteristic size of the hatch area distribution. Secondly, it implies a characteristic past time horizon, beyond which we will not be able to backtrack because the possible starting places cover all possible spawning areas. We will return to this issue in the section ‘Discussion’.

Having now established that reversed time simulation is the appropriate methodology to identify hatch areas/hatch schedule probability distributions from a given larval catch, we will focus on reversed time simulation in the rest of this paper. We also note that probability distributions are obtained from larval ensemble trajectories by any standard smoothing technique. The spatial extent of larval distributions are obtained either from form parameters of the smooth distributions fitting larval ensemble positions, or by identifying areas where probabilities are larger than a given tolerance level. All qualitative conclusions below are unaffected by these technical steps and choices. Formally, Eqs. (1) to (3) cover 2 process classes: deterministic advection processes (by fields G , u , a) and local dispersal $d\Omega$. Both these process classes are unambiguously reversible: deterministic advection terms change sign, whereas local dispersal $d\Omega$ keeps its sign. This is due to the fact that random walk processes are fundamentally time reversible: it is impossible to judge whether the clock runs forward or backward from a random walk trajectory $s(t)$ in a stationary (or slowly varying) diffusivity field; more directly, $s(t)$ and $s(-t)$ are equally likely. Further, since the stationary state of a dispersive process is spatially uniform, the backward dynamics dispersion amplitude corresponds to the forward dis-

persion amplitude K . In other words, $d\Omega$ is formally invariant under time reversal. Hence, each larva in the ensemble is propagated from time t to $t - dt$ by the dynamical equations:

$$dL_i = -G(L_i, s_i, t)dt \quad (6)$$

$$ds_i = -[u(s_i, t) + a(s_i, L_i, t)]dt + d\Omega(s_i, t, dt) \quad (7)$$

$$d\Omega(s_i, t, dt) = \nabla K(s_i, t)dt + \sqrt{2K[s_i + \frac{1}{2}\nabla K(s_i, t)dt, t]}W(dt) \quad (8)$$

where $dt > 0$ is a small time step backward. We note that Eqs. (6) to (8) are the consistent way of running the forward model, Eqs. (1) to (3), backwards. When any aspect of the forward model is changed, the corresponding change must be performed in the reversed time model. In the section ‘Growth stochasticity’ we will discuss the impact and complications on backtracking, arising when G in Eq. (6) is generalized to a stochastic function reflecting life-history stochasticity.

As an implementation remark, we note that reverse time particle tracking is most simply performed offline, i.e. current fields, temperature, turbulent diffusivity, salinity, etc., are taken from a precalculated database, generated by running a hydrodynamical setup forward in time covering the period of interest. It is possible to run a hydrodynamical model backward in time as well, using adjoined primitive equations (Griffin & Thompson 1996). This avoids large amounts of storage data transfer, but may be more CPU intensive, depending on the hydrodynamical resolution. However, many realistic, operational hydrodynamic setups do not offer this advanced feature.

Apart from this, implementation of Eqs. (6) to (8) is a straightforward modification of the implementation of Eqs. (1) to (3): they are solved by trajectory integration backward in time, exactly like Eqs. (1) to (3) are used forward in time.

SPATIAL AND TEMPORAL BACKTRACKING

To illustrate the approach, we perform backtracking of 3 representative samples of larvae caught at different locations in the North Sea in 2001 (LIFECO 2004). The catch samples are summarized in Table 2.

Table 2. Catch data for 3 representative larval samples used for spatial backtracking examples. $\langle \rangle$ indicates spatial average; σ indicates root mean square

Sample	Catch date (2001)	Catch position	Catch statistics		
			$\langle L \rangle$ (mm)	$\sigma(L)$ (mm)	$\langle \text{age} \rangle$ (d)
1	22 April	56.66° N, 6.66° E	18.2	1.6	23
2	27 April	55.68° N, 7.00° E	19.9	1.1	33
3	25 May	57.15° N, 8.01° E	22.0	1.0	32

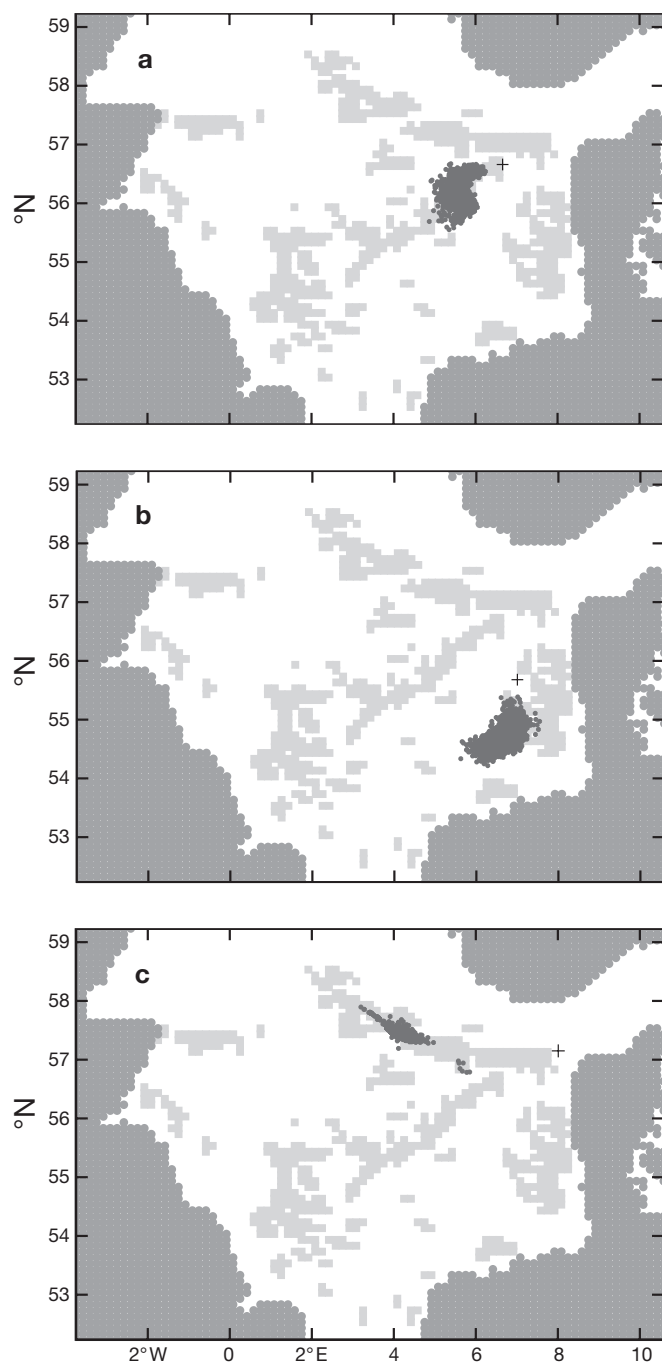


Fig. 4. Backtracking of larval samples in Table 2—(a), (b) and (c) are Samples 1, 2 and 3 respectively—using biological Model v1. Grey: land; light grey: sand banks suitable for sandeel *Ammodytes marinus* habitats; +: larval catch position; small, dark grey circles: backtracked larval ensemble positions that indicate potential hatch positions

In Fig. 4, we show the results of backtracking the larval samples in Table 2. Fig. 4 shows larval catch position, along with the spatial hatch probability distribution. The ensembles have been initialized at the catch position at catch time with normal length distribution,

with form parameters from Table 2, and traced backward, until they have a hatch length of $L = L_0$, using growth Model v1. The figures are overlaid with identified sandeel *Ammodytes marinus* fishing banks, as obtained from detailed fishery loggings (Jensen & Rolev 2004). Habitat data have been projected onto the hydrodynamic grid, so that length-scale features below approximately 10 km are not resolved. It is assumed that larvae must originate from some of these sand banks, as sandeel spawn demersal eggs within their habitats (Reay 1970). Each ensemble size in Fig. 4 contains 2000 individuals, in order to roughly map the spatial hatch probability distribution.

Sample 1 was caught at the eastern tip of the central bank system and was under influence of the cyclonic North Sea circulation system, which, in this area, normally results in northeasterly transport. The sample was traced back to the eastern central bank system around the 'tail end' fishing area (approximately 56° N, 5.5° E), i.e. retained in the same major bank system. Sample 2 was advected along the northward Jutland coastal current, and was traced back to the south-western part of the Jutland sand bank system (approximately 54.5° N, 7° E). Sample 3 was under the influence of Norwegian trench inflow and could be traced back to the middle of the northern major sandeel bank system, around the 'Klondyke' fishing area, approximately 200 km west of catch position.

Comparing the 3 examples, we see prominent differences in the shape and extent of the predicted spatial hatch probability distributions, and also a clear difference in the advection distance, i.e. distance between the most likely hatch position and catch location. This puts clear emphasis on the necessity of explicitly including realistic advection, $u(s,t)$, and dispersal fields, $K(s,t)$, when backtracking larval ensembles. In all considered cases, the hatch probability distribution has a significant overlap with a sand bank system suitable for sandeel habitat (Jensen & Rolev 2004).

We illustrate the sensitivity of backtracked distributions by varying the backtracking model parameters for Sample 2 in Table 2. The central parameters are the horizontal integration time step dt and the growth model coefficients displayed in Table 1. The influence on backtracked larval positions is shown in Fig. 5. Comparing Fig. 5a and b, we see that the effect of changing the horizontal trajectory integration algorithm from Euler to Runge-Kutta 2nd order (Press et al. 1992) (with the same time step) is negligible. If tidal current fluctuation were not averaged out, Euler trajectory integration would display much larger errors for the same time step, due to the rotating nature and high amplitudes of tidal current fluctuation in the North Sea. Comparing Fig. 5a and c, we note the effect of using a quadratic temperature modulation $\lambda(T)$

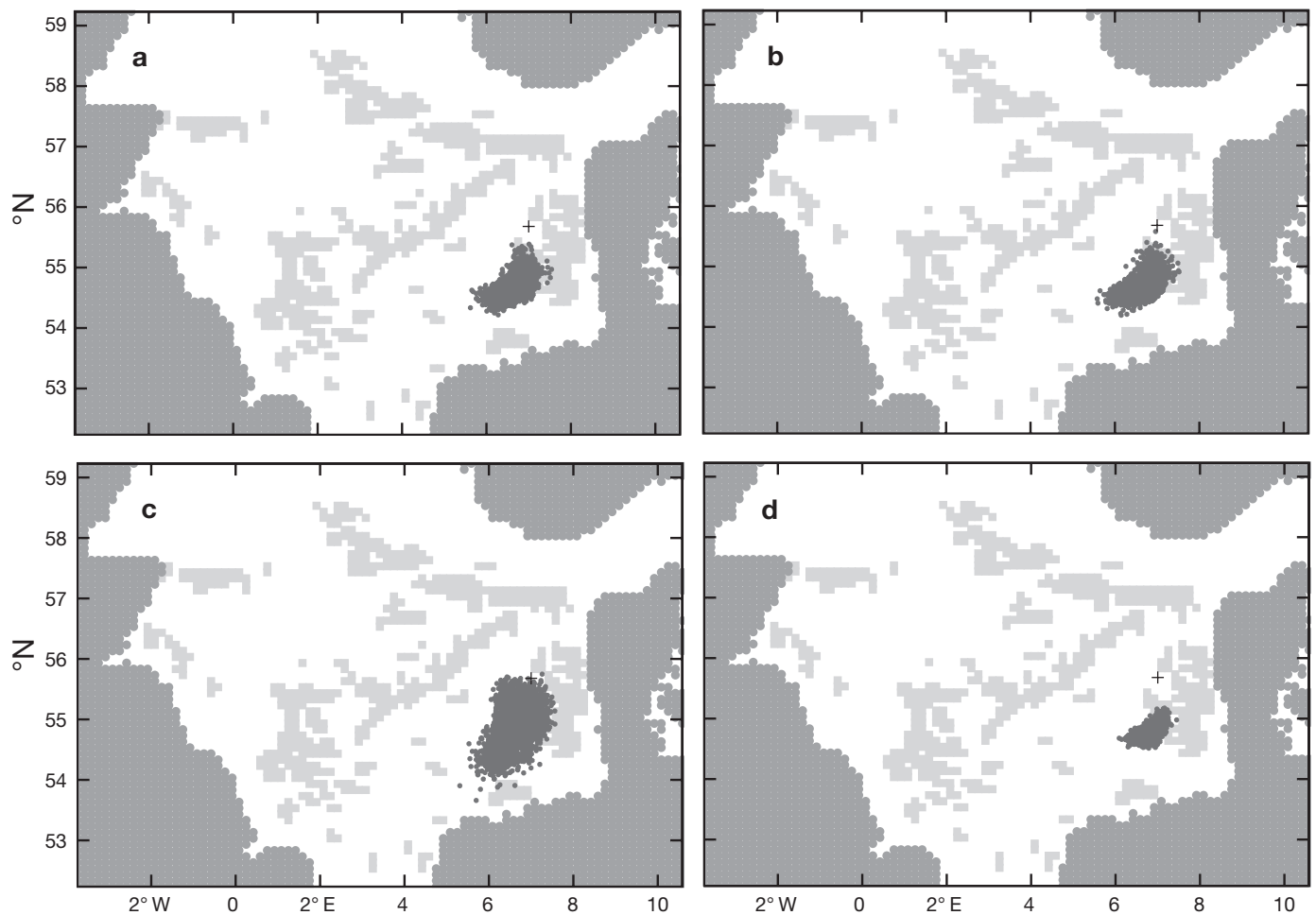


Fig. 5. Sensitivity test on model parameters, shown for backtracking of larval Sample 2 in Table 2. Grey: land; light grey: sand banks suitable for sandeel *Ammodytes marinus* habitats; +: larval catch position; small, dark grey circles: backtracked larval ensemble positions. Biological models/tracing algorithms are (a) v1/Euler, (b) v1/Runge-Kutta 2nd order, (c) v2/Euler and (d) b1/Euler

(Model v2) instead of a linear temperature modulation (Model v1): the predicted hatch areas are similarly centred, but the dispersal is significantly larger. Finally, comparing Fig. 5a and d, we see the effect of using a length-scaling exponent $\gamma = 1$ (Model b1) instead of $\gamma = 0$ (Model v1). Again, the predicted hatch centres are the same, but the dispersal is somewhat smaller. Generally, we predict approximately the same hatch centres for all models, but with some fluctuation in dispersal. There is no systematic bias in dispersal patterns, when comparing models over a larger test set (data not shown). Generally, we find that the effect of changing the horizontal trajectory integration algorithm from Euler to Runge-Kutta 2nd order is negligible.

The representative parameter sensitivity was displayed in Fig. 5; however, under certain hydrographic conditions a larger sensitivity is encountered. The most extreme case in our study is shown in Fig. 6, Sample 3

using Model v1 versus Model b1. In this case, there is an exceptionally large offset between predicted hatch areas, of the order of 100 km, but both are within the same major bank system. This is the largest variation found, and we stress that this case is isolated; the normal variability picture is as illustrated in Fig. 5. However, these isolated cases can also be very useful because they provide a clear prediction for validating specific models, when additional data are present for comparison. Unfortunately, we do not have data to ground truth larvae in Samples 1 to 3 by secondary means.

Until now we have focused on the spatial performance of backtracking. It is also interesting to assess the aptness in the temporal domain. Fig. 7 shows the hatch time distributions from backtracking the larval samples described in Table 2, where the spatial and temporal variability of the local physical environment has been included, i.e. the hatch time distributions for Model v1 correspond to the spatial distributions in

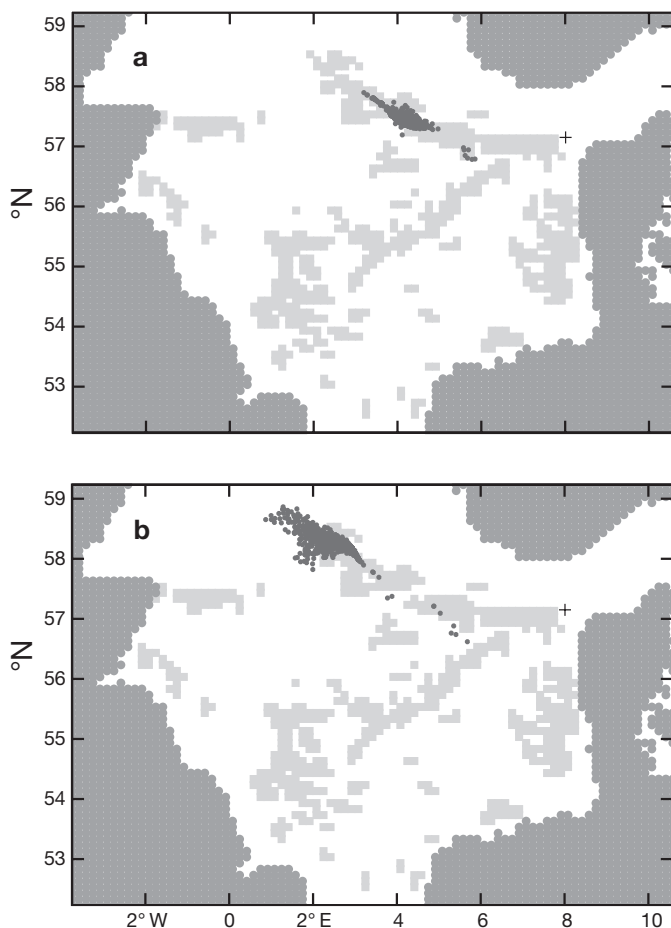


Fig. 6. Anomalous large parameter sensitivity: Sample 3, using (a) Model v1 and (b) Model b1. Grey: land; light grey: sand banks suitable for sandeel *Ammodytes marinus* habitats; +: larval catch position; small, dark grey circles: backtracked larval ensemble positions that indicate potential hatch positions

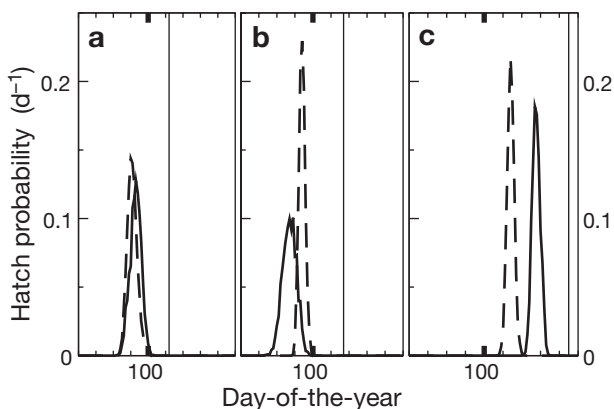


Fig. 7. Hatch time distributions from backtracking of larval samples in Table 2: (a), (b) and (c) are Samples 1, 2 and 3 respectively. Continuous line: Model v1; dashed line: Model b1; thin vertical lines: catch time of sample; tick marks on x-axes: 10 d intervals

Figs. 4 & 5d. Although Models v1 and b1 overlap in the spatial prediction of hatch area, somewhat larger relative differences appear in the prediction of drift period, with Sample 3 again being an outlier, but the overall predicted hatch periods are in reasonable agreement with age assessment from otolith readings (Table 2). Both growth models result in almost the same prediction of hatch peak for Sample 1, Model v1 performs a little better for Sample 2, whereas Model b1 is a little better for Sample 3, so there is no apparent bias toward length of drift period between Models v1 and b1, but both models indicate that there was large spatial heterogeneity in hatch periods over the North Sea in 2001, consistent with survey observations (LIFECO 2004).

GROWTH STOCHASTICITY

Growth variability within an ensemble of larvae arises from many sources, e.g. genetic variability, food patchiness and other environmental fluctuations on a subgrid scale. Growth variability over larger spatial scales will also appear indirectly, if spatial effects are not explicitly represented. From Fig. 2 we can roughly estimate the relative growth variability, i.e. $\sigma(G)/\langle G \rangle$, to be of the order of 20% for sandeel *Ammodytes marinus* larvae in the North Sea (this estimate is an upper limit, because it aliases some spatial and temporal variability as fundamental growth variability, because data in Fig. 2 are pooled). However, it is reasonable to expect different levels of variability for other fish species and other areas.

At this point, it is convenient to shift to a discrete representation in time and larval size, so the larval ensemble is characterized by a distribution vector p_t giving the size distribution of larvae in a suitable set of length classes at time t . If Eq. (1) is integrated forward by a fixed, small time step $dt = h$, the larval length distribution development is characterized by the matrix Γ^h :

$$p_{t+h} = \Gamma^h p_t \quad (9)$$

which can be considered a Markov process, when the time scale of average temperature changes is large compared to h . Uniform mortality can be handled by multiplying a pre-factor to Eq. (9)—this does not alter the qualitative discussion. Put in another way, we focus on the relative length characteristics of a larval ensemble. The matrix Γ_{ij}^h gives the transition probability between length classes j to i during the time step h . Bayes' theorem provides the time-reversed process characterized by the matrix Q^h

$$p_t = Q^h p_{t+h} \quad (10)$$

$$Q^h = (\kappa \cdot \Gamma^h)^T \quad (11)$$

where the matrix κ_{ij} is the ratio of prior probabilities in states j , i and \cdot means element-by-element matrix product. For regular Markov processes (as spatial dispersion), κ is straightforward, but for oriented stochastic processes, like growth, the larval/juvenile length classes have zero probability in the stationary state (all larvae have become adults) and therefore the prior ratio κ is ill defined.

To illustrate the effect of growth stochasticity on larval backtracking, we add a stochastic width of $\sigma(G) = 0.2G$ to the average growth G in Eq. (4), consistent with Fig. 2 (this means absolute growth fluctuations are smaller for early larvae than for juveniles) and again apply the North Sea average temperature variation in the backtracking period. In Fig. 8 we have backtracked growth of larval ensembles for 40 d under different premises. The Markov growth matrix Γ^h in Eq. (11) in this example is obtained by integrating Eq. (4) corresponding to $h = 2$ d and projecting onto size class bins of $\frac{1}{3}$ mm. The backward Markov growth matrix Q^h has then been generated with Eq. (11), using a recently proposed alternative direct scheme (A. Christensen unpubl. data) to compute κ , which resolves the problems with the indefiniteness of κ by the usual definition. Fig. 8a,b shows that growth stochasticity adds width (continuous line) to the sharp length distribution obtained by deterministic backtracking (dashed line) of a narrow length distribution; the length distribution of growth Model b1 (Fig. 8b) is more narrow than that of growth Model v1 (Fig. 8a)

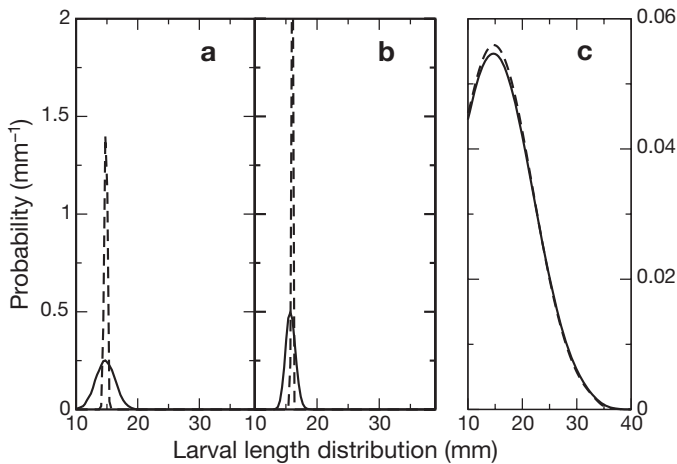


Fig. 8. Larval ensembles backtracked for 40 d, starting on the 120th day of the year, with initial average length $\langle L \rangle = 50$ mm: (a) initial length root mean square (RMS) $\sigma(L) = 0$ mm and growth Model v1, (b) initial length RMS $\sigma(L) = 0$ mm and growth Model b1 and (c) initial length RMS $\sigma(L) = 10$ mm and growth Model v1. Growth models are discretized in size classes of $\frac{1}{3}$ mm, according to Eq. (9). Continuous lines: ensemble backtracking, using Eq. (11), with relative growth stochasticity $\sigma(G) = 0.2\bar{G}$; dashed lines: deterministic ensemble backtracking, using Eq. (6), i.e. $\sigma(G) = 0$

and also biased off centre. We see that Model v1 displays greater sensitivity than Model b1, which can be traced to the fact that Model b has less temperature sensitivity, because length scaling of growth explains the growth variation of very small larvae. Fig. 8c shows that effects of typical growth variability are dominated by typical variability in length distribution. This conclusion, however, need not be true for other species and other areas, especially if backtracking is performed for a longer part of their life history. We also emphasize the importance of a realistic prior κ , if the objective is to recover time scales, e.g. hatch schedules. Finally, we note that a peculiarity of Eq. (11) is that a process which is deterministic in forward time direction may become stochastic in reversed time (or vice versa) due to the transposition of the forward Markov matrix Γ^h .

DISCUSSION

In horizontal dispersal processes, transport distance R and transport time t scales (Taylor 1921) are linked by:

$$R = \sqrt{4\bar{D}t} \quad (12)$$

on a coarse scale, where \bar{D} is a characteristic average effective horizontal dispersal rate. For the North Sea in the larval drift period, Fig. 4 indicates roughly that \bar{D} is $\sim 20 \text{ km}^2 \text{ d}^{-1}$ or $\sim 230 \text{ m}^2 \text{ s}^{-1}$, which is within the normal range, $\bar{D} \approx 100$ to $1000 \text{ m}^2 \text{ s}^{-1}$, of encountered dispersal rates in tidally dominated waters (Zimmerman 1986). This supports the earlier assertion (Zimmerman 1986) that the dominant horizontal dispersion pathway was vertical diffusivity coupled to vertical current shear. Explicit inclusion of other horizontal dispersion pathways can be expected to increase the average effective horizontal dispersal rate \bar{D} somewhat, so our diffusivity fields are lower bounds on the total diffusivity. However, we believe that it will not change our conclusions, nor the qualitative picture we draw in our paper, because they hold also for increased spatial diffusivity fields. In the future, we plan to include other horizontal dispersion pathways for the sake of completeness.

In our context, this relation has 2 important implications

- A time horizon t_c : if we want to backtrack larvae and localize them with a spatial accuracy R (or better), we can only simulate backwards for $t < t_c = R^2 / 4\bar{D}$, before dispersion exceeds the desired spatial accuracy R . It should also be observed that R exceeds 2-fold the spatial grid resolution, which in our case is $2 \times 10 \text{ km}$ (this is the resolution limit of current structures)—otherwise the spatial grid resolution should be increased correspondingly. This lower grid resolu-

tion limit corresponds to roughly 5 d backtracking in the North Sea

- Spatial resolution limit R_c : if we want to simulate backwards for a specific time period t the spatial accuracy of the answer is $R_c = \sqrt{4Dt}$. For lesser sandeels in the North Sea, the drift period is on the order of 2 mo, so that the fundamental backtracking resolution for settled juveniles is roughly $R_c \approx 70$ km. This is somewhat larger than the typical area resolution in Fig. 4, which concerns presettled larvae but includes cohort length variation (see Table 2). On the other hand, including cohort length distribution directly gives the time distribution of a probable hatch.

Spatial backtracking provides an initial probability distribution. However, care must be taken in the interpretation of these probability fields. Batchelder (2006) suggested that the fraction of a planktonic sample originating from a specific area can be determined by considering overlap with the backtracked probability distribution; this can only be accomplished if all sites of origin have a uniform prior distribution of the planktonic organisms in question. Further, the mortality must be spatially uniform. Batchelder (2006) also suggested time round-trip experiments as an appealing and intuitive validation step of backtracking schemes. Here, a given situation is first run forward in time for a selected interval. Then, the resulting spatial distribution is run backward in time for the same time interval, and the final spatial distribution is overlaid with the initial state. In the light of our comments above, we want to emphasize that the final spatial distribution (i.e. after the time round trip) should not generally be expected to be strongly overlapping with the initial state. The overlap can be made arbitrarily small or large, depending on the choice of the round-trip time interval. It is not a question of good or bad performance, but of the speed of information loss, which is determined by the local diffusivity. Only a negative validation is possible, if strictly no overlap exists between final spatial distribution and initial state. We also want to note that the centre of gravity in the final spatial distribution should not be expected to coincide with the centre of gravity in the initial state, when the diffusivity is spatially heterogeneous.

Another interesting mechanism for particle dispersal is the potential presence of Lagrangian chaos (Zimmerman 1986). The presence of Lagrangian chaos is determined completely by the structure of the advective flow field transporting the particles. If it is due to spatial current structures above grid scale, it is automatically picked up by trajectory integration. However, if it is due to subgrid scale spatial current structures, it must appear as a contribution to the spatial turbulent density $K(s,t)$. An interesting aspect in this

context is that diffusive trajectories starting close to each other diverge as the square root of time, whereas chaotic trajectories diverge exponentially with time. Since we have demonstrated the importance of using realistic dispersal fields $K(s,t)$, future studies should address the nature and parameterization of subgrid scale Lagrangian chaos contributions to particle dispersal, as well as disentangling the overlap with standard turbulent dispersal parameterizations, to avoid double accounting.

Even though the hydrodynamic database underlying our study has been validated in detail by available observations (Janssen et al. 2001, Janssen 2002), as pointed out in the section 'Biological and physical model', some level of circulation field uncertainty must be expected to be present. At the most crude level, this can be modelled as an additional, homogeneous diffusivity. In this case, it will not affect the qualitative conclusions of this paper. At a more sophisticated level, this can be modelled as a spatially and temporally dependent autocorrelated random walk process, overlain on the particle tracking described above. However, it is beyond the scope of the present paper to parameterize such an error model process.

We have not yet discussed the choice of larval ensemble size N , because we have focused on qualitative results rather than quantitative results. Two guiding figures apply here. If we want to observe a biological event in our simulations (e.g. a larva crossing a critical point in its life cycle), which has an approximate probability p , then $1/p < N$ should be satisfied. Conversely, if we are limited by computer resources to an ensemble size of N_c , we can only expect to observe events more likely than $1/N_c < p$. If we want to estimate p numerically in our particle tracking (with dispersal), we are bounded by the counting noise, which gives the relative variance of our estimate $\sigma(p)/p \approx 1/\sqrt{pN}$. This criterion is harsh to meet in reality, if p is small, and acceleration techniques, like Brownian Bridges (Rogers & Williams 1987) may become necessary.

Finally, we want to discuss the influence of mortality on our results. We have focused on virtual larval trajectories. If particles are passive and without internal states (e.g. size and condition, or mortality is independent of internal states), these trajectories can be corrected *a posteriori* by any mortality schedule. If mortality is spatially homogeneous, mortality effects will not affect many relative properties (like relative survival of ensemble sub-groups) or survival chance will trivially depend on drift time. In the case of a spawning-site distribution backtracked from a catch location, like in Fig. 4, spatially homogeneous mortality will not change the predicted spawning-site distribution. Thus, one could say that all our examples in the present paper are also valid for a constant mortality level. If the

mortality level is spatially and/or temporally varying, changes in the predicted spawning-site distribution must be expected.

An interesting example to study would be the impact of spatial predator distribution; in this case, survival chances along virtual larval trajectories (both forward and backward in time) can be assessed as path integrals along virtual particle trajectories of the spatial predator distribution. Virtual particle trajectories have another advantage, when comparing different mortality scenarios: the same set of particle trajectories can be computed once, stored, and used for all mortality scenarios. Re-using trajectories has the added advantage that the $1/\sqrt{N}$ sampling noise level is suppressed. If particles with internal states in forward tracking die, they are removed from the ensemble. What happens in backtracking? Then, they must be added to the ensemble at a rate corresponding to the local mortality. The less trivial question is what internal state (i.e. age, condition, etc.) should be assigned to these added particles. Clearly more research is needed to address this open question.

CONCLUSIONS

Forward simulation has often been used as a device to discriminate different potential underlying biological mechanisms. We have demonstrated larval backtracking as a versatile tool, complementary to normal forward simulations for model validation. Backward processes may exhibit a sensitivity not present in the forward processes. Model sensitivity is traditionally considered a weakness, because it makes assumptions important. Conversely, strong model sensitivity can be considered a powerful model validation asset, in the presence of auxiliary data—we advocate the latter point of view, used carefully, as a constructive attitude.

We have illustrated larval backtracking for North Sea lesser sandeel *Ammodytes marinus* larvae and tested several alternative biological growth models. We have found that dispersal effects are important for larval backtracking predictions, with large differences in shapes and extent of predicted hatching areas for larval patches originating from different regions of the North Sea, as well as large differences in the average advection distance. This emphasizes the need for future studies on the quality of sub-grid-scale turbulence parameterization, including possibly the effect of sub-grid-scale Lagrangian chaos. We have found backward prediction of hatch area more robust to model parameters than backward prediction of temporal hatch schedule. In all considered cases, a biologically reasonable hatching area has been predicted, without nudging the model in this direction. We have

generally found reasonable agreement between backward predicted temporal hatch distribution and otolith reading data.

We have clarified fundamental limitations of larval backtracking due to information loss in stochastic processes, most importantly the time horizon and spatial resolution limit for backward hatch area prediction. For juvenile sandeel larvae in the North Sea, the lower backtracking resolution limit is 70 km, increasing if growth stochasticity and/or cohort length variance is included. The accuracy of backtracking is bounded by turbulence processes on long (monthly) time scales and by spatial hydrodynamic resolution on short (weekly) time scales.

Finally, we have found that growth stochasticity adds an uncertainty to the backtracked hatch estimate, similar to that in hatch area prediction. For North Sea lesser sandeel, we estimate the relative growth rate variability to be of the order of 20%. If the larval ensemble has a broad length distribution, it may be sufficient to use deterministic backtracking (i.e. use average growth instead of stochastic growth), provided growth stochasticity is small or moderate, which has been found to be the case for North Sea sandeels.

In a backtracking perspective, the most rewarding development on the biological side will be linking growth variability to the local biophysical environment, in order to capture the effects of regional, seasonal and interannual differences in environmental conditions with respect to growth variability. Also, more work is needed in elaborating unresolved processes and uncertainties on both the biological side, with emphasis on active behaviour and physical cue responses, and on the hydrodynamic side, with emphasis on improving the parameterization of hydrodynamic dispersal fields.

Acknowledgements. This work has, in part, been supported by EU FP6/SSP projects BECAUSE (Contract No. 502482) and PROTECT (Contract No. 513670).

LITERATURE CITED

- Allain G (2004) Biophysical modelling for recruitment prediction. PhD thesis, Ecole Nationale Supérieure Agronomique de Rennes (in French)
- Allain G, Petitgas P, Grellier P, Lazure P (2003) The selection process from larval to juvenile stages of anchovy (*Engraulis encrasicolus*) in the Bay of Biscay investigated by Lagrangian simulations and comparative otolith growth. Fish Oceanogr 12(4–5):407–418
- Arnott SA, Ruxton GD (2002) Sandeel recruitment in the North Sea: demographic, climatic and trophic effects. Mar Ecol Prog Ser 238:199–210
- Baron PR (2004) The life history strategy of sprat (*Sprattus sprattus*) in the Baltic ecosystem. PhD thesis, Universidad de Cadiz

- Batchelder HP (2006) Forward-in-time/backward-in-time-trajectory (fitt/bitt) modeling of particles and organisms in the coastal ocean. *J Atmos Ocean Technol* 23:727–741
- Bergstad OA, Høines AS, Jørgensen T (2002) Growth of sandeel, *Ammodytes marinus*, in the northern North Sea and Norwegian coastal waters. *Fish Res* 56(1):9–23
- Daan N, Bromley PJ, Hislop JRG, Nielsen NA (1990) Ecology of North Sea fish. *Neth J Sea Res* 26(2–4):343–386
- Dechter R, Frost D (2002) Backjump-based backtracking for constraint satisfaction problems. *Art Intellig* 136(2):147–188
- Gallego A, Heath MR, Cook B (2004) The origin and destination of sandeel larvae sampled in the northern North Sea: bio-physical modelling simulation result. ICES CM/P:9
- Griffin DA, Thompson KR (1996) The adjoint method of data assimilation used operationally for shelf circulation. *J Geophys Res* 101:3457–3478
- Hadamard J (1923) Lectures on Cauchy's problem in linear partial differential equations. Yale University Press, New Haven, CT
- Hochbaum U (2004) Modellierung hydrodynamischer Einflüsse auf den Lebenszyklus von *Crangon crangon* in der Nordsee. Master's thesis, University of Hamburg
- Hunter JR, Craig PD, Phillips HE (1993) On the use of random-walk models with spatially-variable diffusivity. *J Comput Phys* 106(2):366–376
- Janssen F (2002) Statistical analysis of multi-year hydrographic variability in the North Sea and Baltic Sea. Validation and correction of systematic errors in a regional ocean model. PhD thesis, Fachbereich Geowissenschaften, Universität Hamburg (in German)
- Janssen F, Schrum C, Huebner U, Backhaus JO (2001) Validation of a decadal simulation with a regional ocean model for North Sea and Baltic Sea. *Clim Res* 18:55–62
- Jensen H (2001) Settlement dynamics in the lesser sandeel *Ammodytes marinus* in the North Sea. PhD thesis, University of Aberdeen
- Jensen H, Rolev AM (2004) The sandeel fishing grounds in the North Sea. Information about the foraging areas of the lesser sandeel *Ammodytes marinus* in the North Sea. Working document prepared for the BECAUSE project. Technical report, Danish Institute of Fisheries Research, Copenhagen
- Jensen H, Wright PJ, Munk P (2003) Vertical distribution of pre-settled sandeel (*Ammodytes marinus*) in the North Sea in relation to size and environmental variables. *J Mar Sci* 60(6):1342–1351
- Kimura S, Kishi MJ, Nakata H, Yamashita Y (1992) A numerical analysis of population dynamics of the sand lance (*Ammodytes personatus*) in the eastern Seto Inland Sea. *Fish Oceanogr* 1:321–332
- Kishi MJ, Kimura S, Nakata H, Yamashita Y (1991) A bio-mass-based model for the sand lance (*Ammodytes personatus*) in Seto Inland Sea. *Jpn Ecol Model* 54:247–263
- Letcher BH, Rice JA, Crowder LB, Rose KA (1996) Variability in survival of larval fish: disentangling components with a generalized individual-based model. *Can J Fish Aquat Sci* 53(4):787–801
- LIFECO (Linking hydrographic Frontal activity to ECOSystem activity) (2004) EU Fifth Framework Programme research project LIFECO (Q5RS-2000-30183), final report. LIFECO, Danish Institute of Fisheries Research, Copenhagen
- Macer CT (1966) Sand eels (*Ammodytidae*) in the south-western North Sea; their biology and fishery. *Fish Invest* 24(6):1–55
- Maier-Reimer E (1973) Hydrodynamisch-numerische Untersuchungen zu horizontalen Ausbreitungs- und Transportvorgängen. Mitt IFM Univ Hamb 21:56
- Nagoshi M, Sano M (1979) Population studies of sand eel, *Ammodytes personatus*, in Ise Bay. I. Growth and its relation to population density. *Jpn J Ecol* 29:1–10
- Pedersen OP, Tande KS, Slagstad D (2000) A synoptic sampling method applied to *Calanus finmarchicus* population on the Norwegian mid-shelf in 1997. *Mar Ecol Prog Ser* 204:143–157
- Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1992) Numerical recipes in C: the art of scientific computing. Cambridge University Press, New York
- Reay PJ (1970) Synopsis of biological data on North Atlantic sandeels of the genus *Ammodytes*. FAO Fish Synop 82, FAO, Rome
- Rogers LCG, Williams D (1987) Diffusions, Markov processes and martingales, Vol II. Ito calculus. Wiley, New York
- Schrum C (1997) Thermohaline stratification and instabilities at tidal mixing fronts. Results of an eddy resolving model for the German bight. *Cont Shelf Res* 17(6):689–716
- Schrum C, Backhaus JO (1999) Sensitivity of atmosphere-ocean heat exchange and heat content in North Sea and Baltic Sea. A comparative assessment. *Tellus* 51A:526–549
- Schrum C, Siegmund F, St John M (2003) Decadal variations in the stratification and circulation patterns of the North Sea. Are the 90's unusual? In: ICES symposium of hydro-biological variability in the ICES area 1990–1999. *J Mar Sci* 219:121–131
- Schrum C, Alekseeva I, St John M (2006) Development of a coupled physical-biological ecosystem model ECOSMO, Part I: model description and validation for the North Sea. *J Mar Syst* 61(1–2):79–99, doi:10.1016/j.jmarsys.2006.01.005
- Smigielski AS, Halavik TA, Buckley LJ, Drew SM, Laurence GC (1984) Spawning, embryo development and growth of the American sand lance *Ammodytes americanus* in the laboratory. *Mar Ecol Prog Ser* 14:287–292
- Spivakovskaya D, Heemink AW, Milstein GN, Schoenmakers JGN (2005) Simulation of the transport of particles in coastal waters using forward and reverse time diffusion. *Adv Wat Res* 28(9):927–938
- Taylor GI (1921) Diffusion by continuous movements. *Proc Lond Math Soc* 20:196–211
- Uliasz M, Pielke RA (1991) Application of the receptor oriented approach in mesoscale dispersion modeling. In: van Dop H, Steyn DG (eds) Air pollution modeling and its application. VIII. Plenum Press, New York, p 399–408
- Van Dam GC, Ozmidov RV, Korotenko KA, Suijlen JM (1999) Spectral structure of horizontal water movement in shallow seas with special reference to the North Sea, as related to the dispersion of dissolved matter. *J Mar Syst* 21(1–4):207–228
- Visser AW (1997) Using random walk models to simulate the vertical distribution of particles in a turbulent water column. *Mar Ecol Prog Ser* 158:275–281
- Winslade P (1971) Behavioral and embryological studies on the lesser sandeel *Ammodytes marinus* (Raitt). PhD thesis, University of East Anglia, Norwich
- Wright PJ, Bailey MC (1996) Timing of hatching in *Ammodytes marinus* from Shetland waters and its significance to early growth and survivorship. *Mar Biol* 126(1):143–152
- Zimmerman JTF (1986) The tidal whirlpool—a review of horizontal dispersion by tidal and residual currents. *Neth J Sea Res* 20(2–3):133–154